Little Mathematics Library



A.I. MARKUSHEVICH

REMARKABLE CURVES

Mir Publishers · Moscow

ПОПУЛЯРНЫЕ ЛЕКЦИИ ПО МАТЕМАТИКЕ

А. И. Маркушевич

ЗАМЕЧАТЕЛЬНЫЕ КРИВЫЕ

ИЗДАТЕЛЬСТВО «НАУКА» МОСКВА

LITTLE MATHEMATICS LIBRARY

A.I. Markushevich

REMARKABLE CURVES

Translated from the Russian
by
Yu. A. Zdorovov

MIR PUBLISHERS
MOSCOW

First published 1980
Revised from the 1978 Russian edition

На английском языке

CONTENTS

PREFACE
TO THE THIRD RUSSIAN EDITION

This book has been written mainly for high school students, but it will also be helpful to anyone studying on their own whose mathematical education is confined to high school mathematics. The book is based on a lecture I gave to Moscow schoolchildren of grades 7 and 8 (13 and 14 years old).

In preparing the lecture for publication I expanded the material, while at the same time trying not to make the treatment any less accessible. The most substantial addition is Section 13 on the ellipse, hyperbola and parabola viewed as conic sections.

For the sake of brevity most of the results on curves are given without proof, although in many cases their proofs could have been given in a form that readers could understand.

The third Russian edition is enlarged by including the results on Pascal's and Brianchon's theorems (on inscribed and circumscribed hexagons), the spiral of Archimedes, the catenary, the logarithmic spiral and the involute of a circle.

A. I. Markushevich

1. The Path Traced Out by a Moving Point

In the spoken language the adjective "curved" describes something which continuously deviates from a straight line: a curved form, outline, thing and even fire.

Mathematicians use the word "curve" in the sense of "a curved line". But what is a curve? How can one embrace in a single notion all the curves that are traced out on paper with a pencil or a pen, on a blackboard with a piece of chalk, or in the night sky by a "shooting star" or a rocket?

We shall use the following definition: a curve is the path traced out by a moving point. In our examples the role of the point is played by a pencil point, the sharp edge of a piece of chalk, a burning meteor passing through the upper levels of the atmosphere, or a rocket. According to this definition a straight line is just a particular curve. Indeed, why should not a moving point trace out a straight path?

2. The Straight Line and the Circle

A moving point describes a straight line when it passes from one position to another along the shortest possible path. A straight line can be drawn with the help of a ruler; when a pencil runs along the edge of a ruler it leaves a trace on the paper in the form of a straight line.

When a point moves on a surface at a constant distance from another fixed point on the same surface it describes a circle. Because of this property of the circle we are able to draw a circle with the help of compasses.

The straight line and the circle are the simplest and at the same time the most remarkable curves as far as their properties are concerned. You are no doubt more familiar with these two curves than with others. But you should not imagine that you know all of the most important properties of straight lines and curves. For example, you may not know that if the vertices of the triangles ABC and $A'B'C'$ lie on three straight lines intersecting at the point S (Fig. 1), the three points of intersection M, K, L of the corresponding sides of the triangles, the sides AB and $A'B'$, BC and $B'C'$, and AC and $A'C'$, must be collinear, that is, they lie on a single straight line.

You are sure to know that a point M moving in a plane equidistantly from two fixed points, say F_1 and F_2, of the same plane, that is,

so that $MF_1 = MF_2$, describes a straight line (Fig. 2). But you might find it difficult to answer the question: What type of curve will point M describe if the distance of M from F_1 is a certain number of times

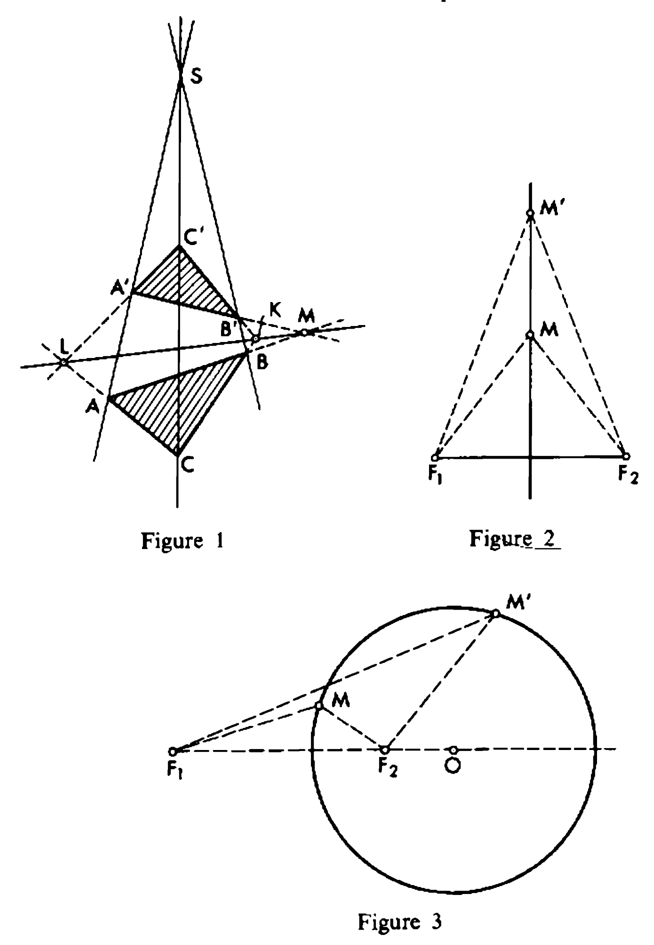


Figure 1

Figure 2

Figure 3

greater than that from F_2 (for instance, in Fig. 3 it is twice as great)? The curve turns out to be a circle.

Hence if the point M moves in a plane so that the distance of M from one of the fixed points, F_1 or F_2, in the same plane is always proportional to the distance from the other fixed point, that is

$$MF_1 = k \times MF_2$$

then M describes either a straight line (when the factor of proportionality is unity) or a circle (when the factor of proportionality is other than unity).

3. The Ellipse

Consider a curve described by a point M so that the sum of the distances of M from two fixed points F_1 and F_2 is constant. Take a piece of string and tie its ends to two pins stuck into a sheet of paper, leaving it loose for a while. If we make the string taut with a vertical pencil and then move the pencil (Fig. 4), the pencil point, M, will describe an oval-shaped curve which looks like a flattened circle; it is called an ellipse.

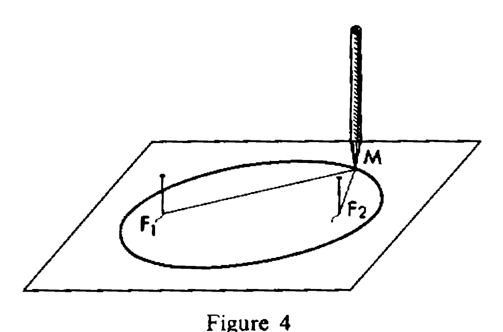


Figure 4

In order to get a closed ellipse we have to move the string over to the other side of the pins after completing the first half. It is obvious that the sum of the distances of the point M from the pin-holes F_1 and F_2 is constant throughout the movement and is equal to the length of the string.

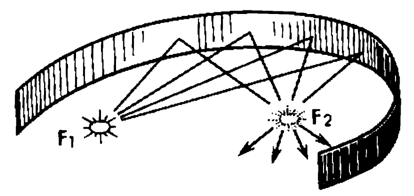


Figure 5

The pin-holes mark two points on the paper called the foci of the ellipse. The word "focus" in Latin means "hearth" or "fire" which is justified by the following remarkable property of the ellipse.

If we construct an ellipse from a strip of polished metal and place some point source of light at one focus, the rays of light emanating from this focus will converge at the other focus. Thus we can see the "fire" (the image of the first focus at the second focus (Fig. 5)).

4. The Foci of an Ellipse

If we draw a straight line through the foci of an ellipse and extend it in both directions till it intersects the ellipse, we shall get the major axis of the ellipse, A_1A_2 (Fig. 6). The ellipse is symmetric with respect to its major axis. By erecting a perpendicular to the line segment F_1F_2 at its centre and extending it till it intersects the ellipse we obtain the minor axis, B_1B_2, which is also an axis of symmetry of the ellipse. The ends A_1, A_2, B_1 and B_2 of the axes are called the vertices of the ellipse.

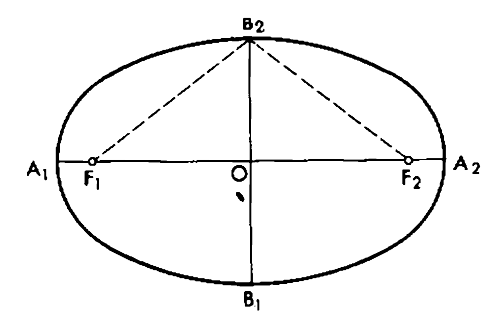


Figure 6

The sum of the distances of the point A_1 from the foci F_1 and F_2 must be equal to the length of the string:

$$A_1F_1 + A_1F_2 = l$$

But because of the symmetry of the ellipse we have

$$A_1F_1 = A_2F_2$$

Hence A_1F_1 can be replaced by A_2F_2 so that

$$A_2F_2 + A_1F_2 = l$$

It is easily seen that the quantity on the left hand side of the equation is the length of the major axis of the ellipse. Thus the length of the major axis of the ellipse is equal to the length of the string. In other words, the sum of the distances to any point of the ellipse from the foci is equal to the length of the major axis of the ellipse. Consequently, from the symmetry of the ellipse we conclude that the distance from the vertex B_2 (or B_1) to a focus is equal to half the length of the major axis. Hence knowing the vertices of the ellipse, we can easily construct its foci. To do so we have to mark off the major axis by the arc of the circle with B_2 as centre and with the radius equal to half of A_1A_2.

5. The Ellipse is a Flattened Circle

Using the major axis of the ellipse as a diameter, we construct a circle (Fig. 7). Through a point N on the circumference erect the perpendicular NP to the major axis. The perpendicular intersects the ellipse at a point M. It is evident that NP is several times greater than MP. It

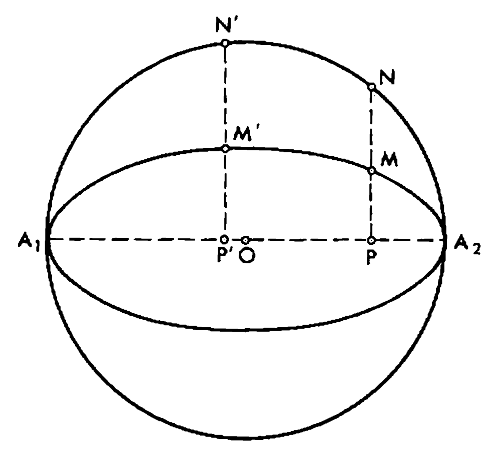


Figure 7

turns out that if we take some other point N' on the circumference and repeat the construction, $N'P'$ will be the same number of times longer than the corresponding line segment $M'P'$, that is

$$NP/MP = N'P'/M'P'$$

In other words, the ellipse can be constructed with the help of the circumscribed circle by making all the points of the circumference proportionally closer to the major axis. A very simple method of constructing an ellipse is based on this property. By drawing a circle and one of its diameters and replacing the points of the circumference by points on the perpendiculars to the diameter that are several times (2, 3 and so on) closer to it we obtain points of an ellipse whose major axis coincides with the diameter and whose minor axis is the same number of times (2, 3 and so on) smaller than the diameter.

6. Ellipses in Everyday Life and in Nature

In everyday life we often come across ellipses. If, for instance, we tilt a glass of water, the edge of the water surface forms an ellipse (Fig. 8). Similarly, if we cut a slice of a cylindrical sausage with the knife at an oblique angle, the slice will have the shape of an ellipse (Fig. 9). In

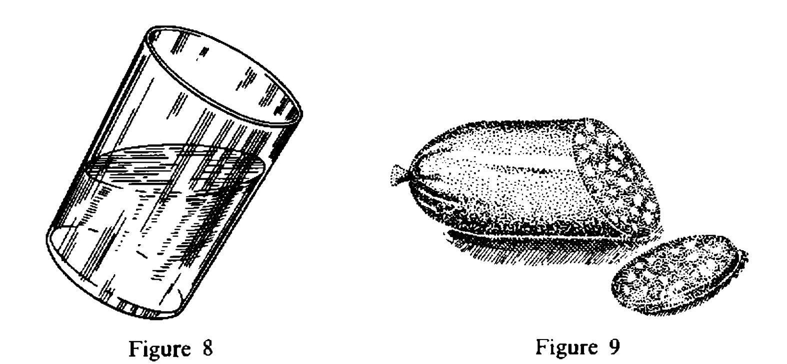


Figure 8 Figure 9

general, if we cut through a right cylinder (or a cone) at an angle without cutting the base, the cut-away view will be an ellipse (Fig. 10).

Johannes Kepler (1571-1630) discovered that the planets move about the sun not along circles, as it had been thought before, but along ellipses with the sun at a focus (Fig. 11). During each period of revolution the planet passes once through the vertex A_1 closest to the

sun, which is called the perihelion, and once through the vertex A_2 farthest from the sun, the aphelion. The Earth, for example, is at peri-

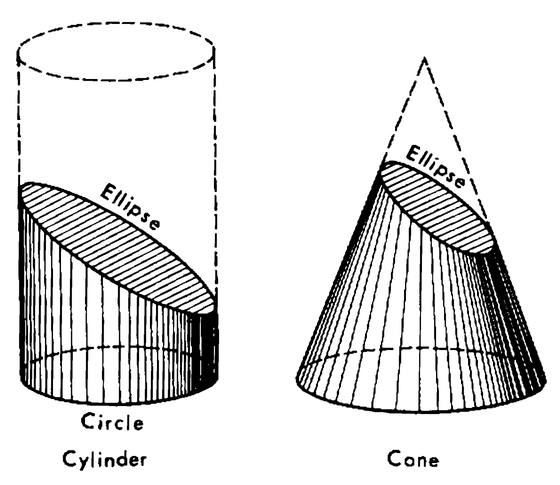


Figure 10

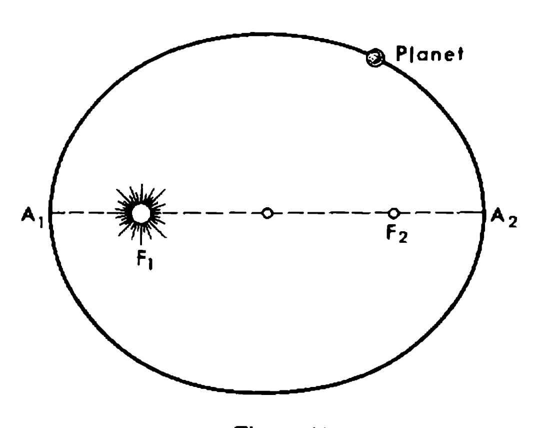


Figure 11

helion when it is winter in our hemisphere and at aphelion when it is summer. The ellipse along which our Earth is moving is flattened only slightly and resembles a circle very closely.

straight line D_1D_2 (called the directrix of the parabola) through the focus is the axis of symmetry of the parabola; usually it is simply called the axis of the parabola.

8. The Parabolic Mirror

If the parabola is constructed from a polished strip of metal, rays from a point source of light at the focus will be reflected parallel to the axis of the parabola (Fig. 13). Conversely, rays parallel to the axis of the parabola will be reflected and brought together at the focus.

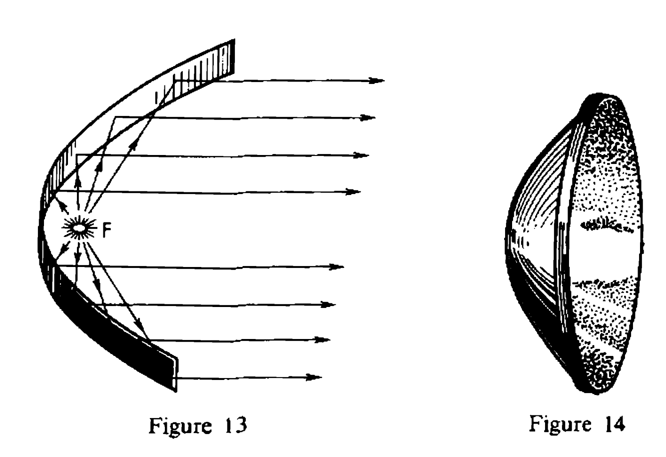


Figure 13 Figure 14

Parabolic mirrors used in automobile headlights (Fig. 14) and in all search lights are constructed in accordance with this property of the parabola. However, they are not made in the form of strips but are paraboloids of revolution. The surface of such mirrors can be obtained by rotating a parabola about its axis.

9. The Flight of a Stone and a Projectile

A stone thrown at an angle (not vertically) travels along a parabola (Fig. 15). This is also true for a projectile such as a cannon ball. In practice air resistance changes the shape of the curve in both cases, so that the resulting curve differs from the parabola. But if we were to observe the motion in a vacuum, we would get a true parabola. By keeping the initial speed of the projectile constant and varying the angle that the projectile makes with the horizontal plane when it

leaves the barrel we shall obtain different parabolas described by the projectile and different ranges. The greatest range will be obtained when the angle of the barrel's inclination is equal to 45° This distance

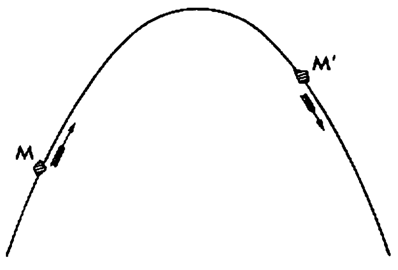


Figure 15

is equal to v^2/g, where g is the acceleration of gravity. When discharged vertically the projectile reaches a height which is half of the greatest range, $v^2/2g$. Irrespective of the angle of the barrel's inclination (keeping the barrel in the same vertical plane), for a given initial

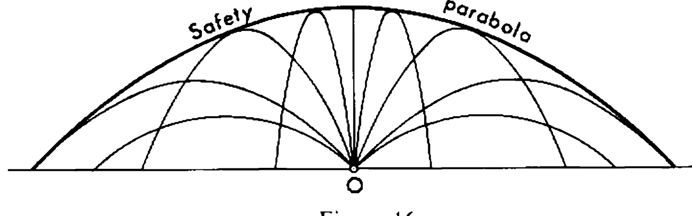


Figure 16

velocity of the projectile there are always places on the ground and in the air where the projectile cannot reach. It turns out that such places are separated from those reachable for some angle of the barrel's inclination by a curve which is also a parabola (Fig. 16), called the safety parabola.

10. The Hyperbola

We can construct curves described by a point M in a way similar to the way we generated the ellipse, only we keep constant not the sum but the difference of the distances from two fixed points F_1 and F_2 or their product or their quotient (in this last case we get a circle). We shall consider now the case when the difference is constant. In order to

make the pencil move in the required way, we fix two pins at the points F_1 and F_2 and rest a ruler against one of the points so that it can rotate on the paper about this pin (Fig. 17). Fix one end of a piece

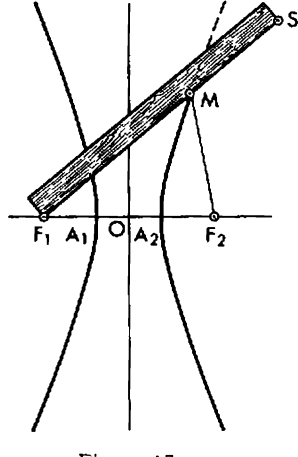


Figure 17

of string to the end S of the ruler (the string should be shorter than the ruler). The other end of the string should be fixed at F_2. Then stretch the string with the help of the pencil point, pressing it against the ruler. The difference of the distances MF_1 and MF_2 is

$$(MF_1 + MS) - (MF_2 + MS) = F_1S - (MF_2 + MS)$$

That is, it is equal to the difference between the length of the ruler and the length of the string. If we rotate the ruler about F_1, pressing the pencil to it and stretching the string, the pencil will describe a curve on the paper for every point of which the difference of the distances from F_1 and F_2 will be the same and equal to the difference m between the length of the ruler and the length of the string. In this way we shall get the upper half of the curve drawn on the right side of Fig. 17. To get the lower half of the curve we have to fix the ruler so that it is below and not above the pins. And finally, if we fix the ruler to the pin F_2 and the end of the string to the pin F_1, we shall get the curve drawn on the left side of Fig. 17. The pair of curves drawn is considered to be one curve which is called a hyperbola; the points F_1 and F_2 are its foci. However, the arcs of the curve depicted do not exhaust all the hyper-

bola. By taking a longer ruler and a longer string (and keeping the difference the same) we can continue our hyperbola indefinitely just as we can continue a straight line segment indefinitely.

11. The Axes and Asymptotes of the Hyperbola

Draw a straight line through the foci of a hyperbola. The straight line is the axis of symmetry of the hyperbola. There is another axis of symmetry perpendicular to the first one bisecting the segment F_1F_2. The point of intersection O is the centre of symmetry and is called the centre of the hyperbola. The first axis cuts the hyperbola at two points A_1 and A_2, called the vertices of the hyperbola; the segment A_1A_2 is called the transverse (real) axis of the hyperbola. The difference between the distances of the point A_1 from the foci F_2 and F_1 must be m:

$$A_1F_2 - A_1F_1 = m$$

But

$$A_1F_1 = A_2F_2$$

because of the symmetry of the hyperbola. Hence A_1F_1 may be replaced by A_2F_2 and we get

$$A_1F_2 - A_2F_2 = m$$

Clearly the difference $A_1F_2 - A_2F_2$ is equal to A_1A_2, that is, to the length of the transverse axis of the hyperbola. So, the difference m of the distances of any point of the hyperbola from the foci (the smaller distance must be subtracted from the greater) is equal to the length of the transverse axis.

From the vertex A_1(or A_2) draw an arc of the circle with radius equal to half of F_1F_2 through the second axis of symmetry. This gives two points B_1 and B_2 (Fig. 18). The segment B_1B_2 is called the conjugate axis of the hyperbola. Then construct the rectangle $PQRS$ whose sides are parallel to the axes of the hyperbola and pass through the points A_1, A_2, B_1 and B_2. Draw its diagonals PR and QS. Extending them indefinitely get two straight lines called the asymptotes of the hyperbola. The asymptotes have an interesting property: they never

intersect the hyperbola, though the points of the hyperbola approach arbitrarily close to the asymptotes, getting closer and closer to them the greater the distance from the centre of the hyperbola. Arcs of the hyperbola lying between points at a considerable distance from the centre resemble a straight line segment in a diagram (see the arc M_1M_2 in Fig. 18), though of course they are nowhere rectilinear; their bending is so slight that it is hardly noticeable.

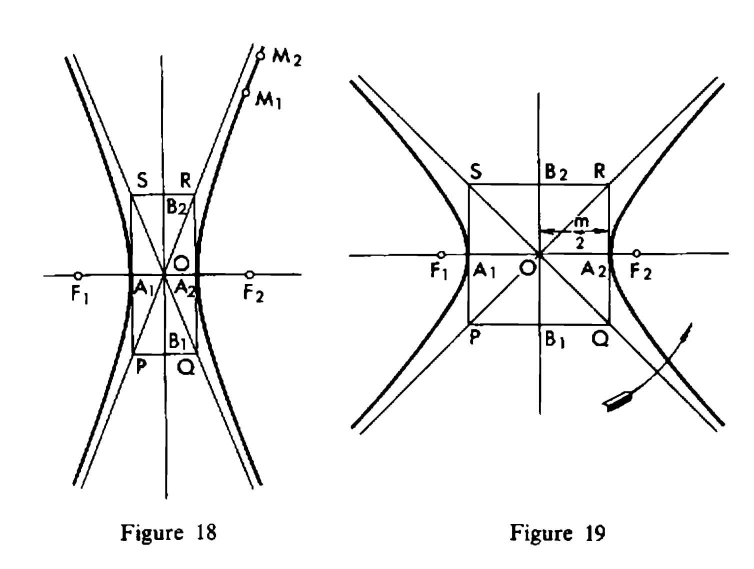


Figure 18

Figure 19

In order to draw a hyperbola in an approximate manner without using a ruler and a string one should do the following. First draw the axes of symmetry of the hyperbola, then mark off two foci F_1 and F_2 on the first of them at an equal distance from the centre, and then, using the same axis, mark off in two opposite directions from the centre segments equal to half of m, that is, half the predetermined difference of the distances of the points of the hyperbola from its foci. This gives us the vertices A_1 and A_2 of the hyperbola. Then mark off the points B_1 and B_2 (as above) on the second axis, construct the rectangular $PQRS$, and finally draw the diagonals. As a result we get a figure shown in Fig. 19. Now only one thing is left to do—to draw two arcs symmetric with respect to the axes and passing through the points A_1 and A_2, bending smoothly and approaching closer and closer the asymptotes PR and QS.

12. The Equilateral Hyperbola

As a particular case the rectangular $PQRS$ can be a square. This is possible if and only if the asymptotes of the hyperbola are mutually perpendicular. In this case the hyperbola is said to be equilateral or rectangular. This particular case is depicted in Fig. 19. For convenience we may rotate the whole figure about the point O through an angle of 45° in the direction indicated by the arrow. In this way we get the hyperbola depicted in Fig. 20. Mark off the segment $ON = x$ on

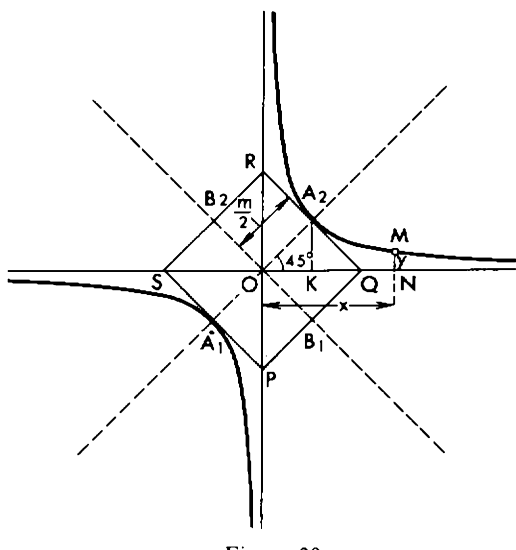


Figure 20

the asymptote OQ, erect the perpendicular $NM = y$ to it at the point N and extend it till it intersects the hyperbola.

There exists a simple dependence between y and x: if x is magnified several times, y is lessened proportionally. Similarly, if x is lessened several times, y is magnified accordingly. In other words, the length of $NM = y$ is inversely proportional to the length of $ON = x$:

$$y = k/x$$

Due to this property the equilateral hyperbola is the graph of inverse proportionality. To find out how the factor of inverse proportionality

k is related to the dimensions of the hyperbola consider the vertex A_2. For this vertex

$$x = OK, \ y = KA_2$$

The line segments OK and KA_2 are the legs of the equilateral right triangle with a hypotenuse

$$OA_2 = m/2$$

Hence,

$$x = y \text{ and } x^2 + y^2 = (m/2)^2 = m^2/4$$

which yields

$$2x^2 = m^2/4 \text{ or } x^2 = m^2/8$$

On the other hand, the equation of the inverse proportionality, $y = k/x$, allows us to conclude that $xy = k$, or in this particular case (where $y = x$) $x^2 = k$. Comparing the two results, $x^2 = m^2/8$ and $x^2 = k$, we find that

$$k = m^2/8$$

In other words, the factor of inverse proportionality k is equal to one-eighth of the square of the length of the transverse axis of the hyperbola.

13. Conic Sections

We have already said that if we cut a cone with a knife, that is, geometrically speaking, intersect it by a plane which does not cut the base of the cone, the boundary (edge) of the section will be an ellipse (Fig. 10). By cutting the cone with a plane so that the cut passes through the base of the cone, we get an arc of a parabola in the cross-section (Fig. 21(*a*)) or an arc of a hyperbola (Fig. 21(*b*)). Thus, all the three curves — the ellipse, the hyperbola and the parabola — are conic sections.

The cone which we used has one drawback: only an ellipse may be placed on it in its entirety (Fig. 10), but a parabola and a hyperbola

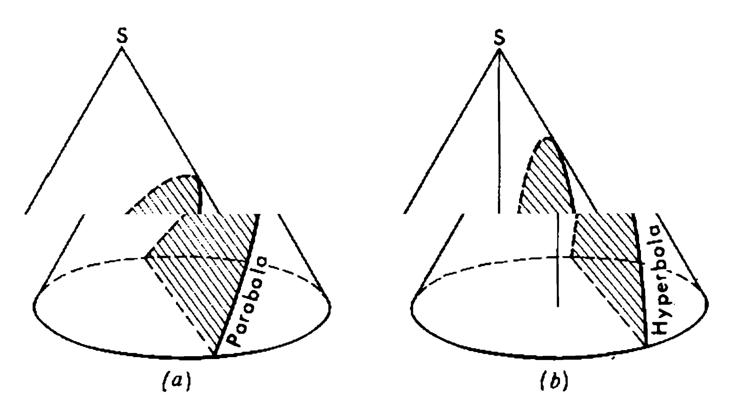


Figure 21

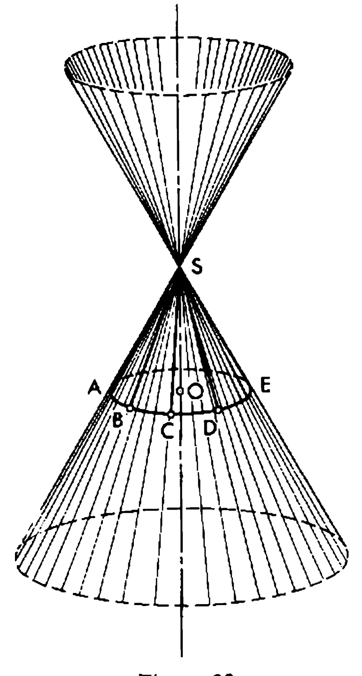


Figure 22

being infinite curves can be incorporated as conic sections only partially. In Fig. 21(*b*) one cannot even see where the second branch of the hyperbola might be taken from. To eliminate this drawback we can replace the cone by an infinite conic surface. To do so we extend all the generators of the cone infinitely far in both directions, that is, we extend the straight line segments *AS*, *BS*, *CS*, *DS*, *ES* and so on connecting the points of the circle at the base of the cone with its vertex (Fig. 22; of course we cannot show the infinite extension of the generators in our figure, so the figure just shows line segments which are longer than the initial ones). The resulting figure will be the required conic surface consisting of two halves connected at the point *S* and extending to infinity. The conic surface can be regarded as the trace of a moving straight line, namely the straight line passing through the point *S* and rotating so that its angle with *OS* (the axis of the conical surface) remains constant. The moving straight line is called a generator of the conical surface; it is evident that the extension of every generator of the initial cone gives a generator of the conical surface.

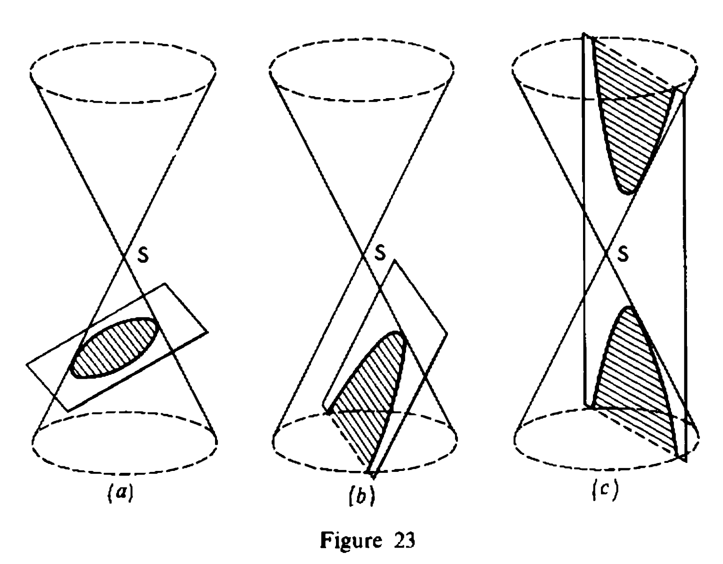


Figure 23

Now let us cut the entire conical surface by a plane. If the plane cuts all the generators in one half of the surface, the plane section is an ellipse (in Fig. 23(*a*) circle); if it cuts all the generators but one parallel to itself, the plane section is a parabola (Fig. 23(*b*)), finally, if the plane

cuts some of the generators in one half of the surface and the remaining generators in the other half, the section is a hyperbola (Fig. 23(*c*)). We can see that one half of the conic surface is enough for an ellipse and parabola. For a hyperbola one needs the whole of the conical surface: one branch of a hyperbola is in one half and the other branch in the other half.

14. Pascal's Theorem

Blaise Pascal (1623-1662) was not yet 17 when he discovered a remarkable property of conic sections. Mathematicians were informed about his discovery by wall posters (50 copies were published but only two of these have survived). Several of these posters were stuck up on the walls of houses and churches in Paris. The reader should not be surprised. At that time (in 1640) there were no scientific journals to inform scientists about new discoveries. Such journals appeared only a quarter of a century later in France and England almost simultaneously. But let us return to Pascal.

Though his advertisement was written in French and not in Latin, as was customary at that time, Parisians gazing at it could hardly make out what it was about. The young genius worded everything extremely concisely and without proof or explanations.

At the beginning of the advertisement after three definitions the author gave "Lemma 1", a theorem which we shall give in different wording. Mark six points on a circle, number them arbitrarily (not

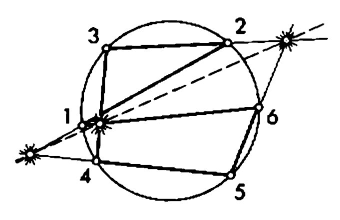


Figure 24

necessarily in their order on the circle) and join them in order by line segments, the last of them connecting point 6 with point 1 (Fig. 24). Pascal's theorem states that the three points of intersection of the straight lines obtained by extending the six line segments and taking them in pairs as follows: the first with the fourth, the second with the fifth and the third with the sixth, all lie on a single straight line (i. e., are collinear).

Try to do this several times yourself taking different points on the circumference (Fig. 25). There may appear a case in which a pair of straight lines whose intersection we are seeking, for example, the first

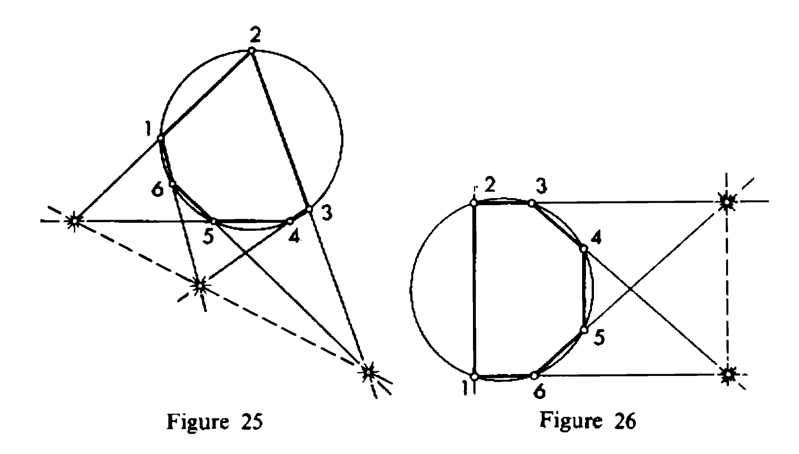


Figure 25

Figure 26

and the fourth, are parallel. In this case Pascal's theorem should be understood as stating that the straight line connecting the two other points of intersection is parallel to the above-mentioned straight lines (Fig. 26). Finally, if in addition to the first and the fourth, the second and the fifth lines are parallel, then in this special case Pascal's theorem states that the straight lines of the last pair, the third and the sixth, are also parallel. We encounter such a case when, for example, the points taken on the circle are the vertices of an inscribed regular

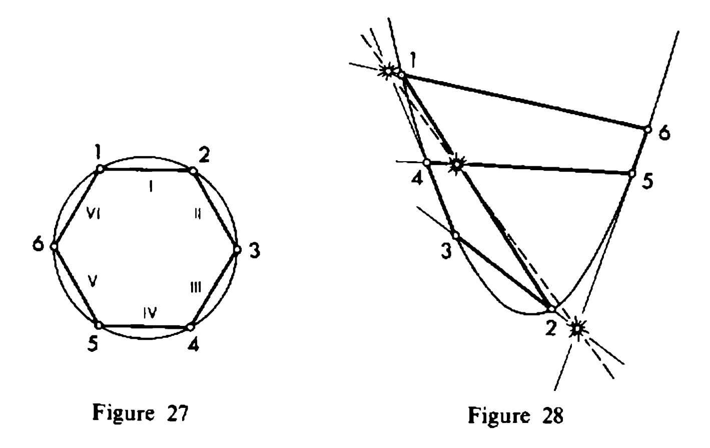


Figure 27

Figure 28

hexagon and are numbered in the order of their position on the circle (Fig. 27).

Pascal did not formulate his theorem only for a circle. He noticed that it held true in the case of any conic section: an ellipse, parabola or hyperbola. Figure 28 illustrates Pascal's theorem for the case of parabola.

15. Brianchon's Theorem

The French mathematician Charles Brianchon (1783-1864) discovered in 1806 that the following theorem, which is the dual of Pascal's theorem, holds true.

Draw six tangents to a circle (or any conic section), number them in an arbitrary order and find the corresponding points of intersection (Fig. 29). Brianchon's theorem states that the three straight lines joining the six points of intersection in pairs as follows: the first with the fourth, the second with the fifth, and the third with the sixth, intersect at a single point (i. e., are concurrent).

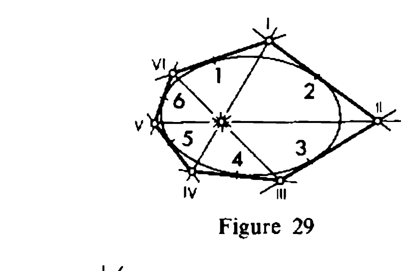


Figure 29

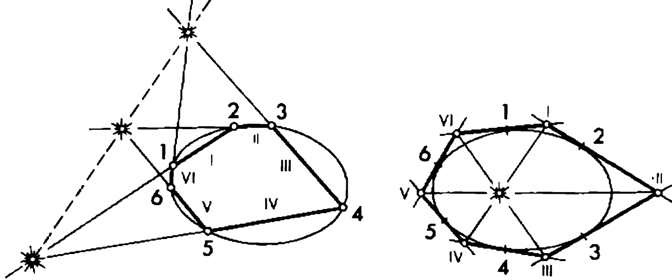


Figure 30

To stress how closely the two theorems are related Brianchon put down both in two columns, one opposite the other (follow Fig. 30

where the left side illustrates Pascal's theorem, and the right that of Brianchon):

Pascal's theorem

Let *1, 2, 3, 4, 5* and *6* be six arbitrary points on a conic section.

Connect them in succession by the straight lines *I, II, III, IV, V* and *VI* and find the three points of intersection of the six lines taken in pairs: *I* with *IV*, *II* with *V* and *III* with *VI*.

Then these three points are collinear.

Brianchon's theorem

Let *1, 2, 3, 4, 5* and *6* be six arbitrary tangents to a conic section.

Find in succession their points of intersection *I, II, III, IV, V* and *VI* and connect them by straight line segments in pairs: *I* with *IV*, *II* with *V* and *III* with *VI*.

Then these lines are concurrent.

It is clear that in order to change from one theorem to the other it is sufficient to make the following interchanges of words and phrases: to interchange "points" and "tangents", "to connect points by line segments" and "to determine the points of intersection of the straight lines", "three points are collinear" and "three straight lines are concurrent". In short, straight lines and points interchange their roles in this transition. In projective geometry conditions are found under which from one true theorem (not necessarily Pascal's theorem) we can obtain another theorem by means of similar interchanges. This is called the principle of duality. It enables us to obtain two theorems for each theorem proved. The other theorem is true, one could say, automatically.

16. The Lemniscate of Bernoulli

Now we shall study a curve generated by a point M in a plane that moves so that the product p of its distances from two fixed points F_1 and F_2 in the same plane is constant. Such a curve is called a lemniscate ("lemniscate" means "with hanging ribbons" in Latin). If the length of the segment F_1F_2 is c, then the distances from the midpoint O of the segment F_1F_2 to F_1 and F_2 are equal to $c/2$ and the product of these distances to $c^2/4$. Let us require at the start that the constant product p should be equal to $c^2/4$, that is, $MF_1 \times MF_2 = c^2/4$. Then the point O will lie on the lemniscate and the lemniscate will have a figure-eight shaped curve lying on its side (Fig. 31). If we extend the segment F_1F_2 in both directions till it intersects the lemniscate, we

shall get two points A_1 and A_2. It is not difficult to express the distance between them, $A_1A_2 = x$, using the known distance $F_1F_2 = c$. Note that the distance from the point A_2 to F_2 is equal to $x/2 - c/2$ and the distance from the same point A_2 to F_1 is equal to $x/2 + c/2$.

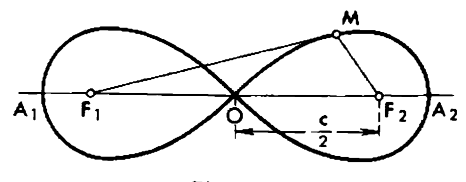


Figure 31

Consequently, the product of the distances is

$$(x/2 + c/2)(x/2 - c/2) = x^2/4 - c^2/4$$

But this product must be equal to $c^2/4$ by assumption. Therefore $x^2/4 - c^2/4 = c^2/4$, which yields $x^2 = 2c^2$ and $x = \sqrt{2}c \approx 1.414c$.

There exists a remarkable relation between this lemniscate and the equilateral hyperbola. Draw rays from the point O (Fig. 32) and mark

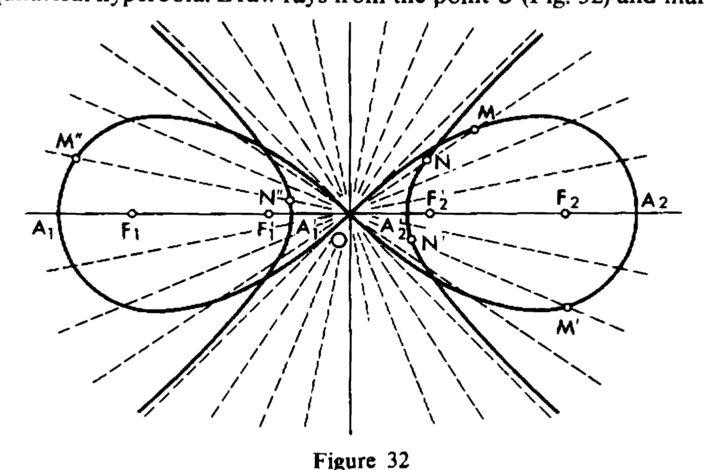


Figure 32

their points of intersection with the lemniscate. It can be seen that when the angle of inclination of a ray to OF_2 (or to OF_1) is less than 45°, the ray intersects the lemniscate at another point distinct from O. But if the angle of inclination is equal to or greater than 45°, there is

no other point of intersection. Take a ray of the first group and assume that it intersects the lemniscate at a point M (distinct from O). Mark off the segment $ON = 1/OM$ from the point O on the ray. If we carry out similar constructions for every ray of the first group, then all the points N corresponding to the points M of the lemniscate will be on an equilateral hyperbola having foci F'_1 and F'_2 such that $OF'_1 = 1/OF_1$ and $OF'_2 = 1/OF_2$.

17. The Lemniscate with Two Foci

If we equate the value of the constant product not to $c^2/4$ but to another value, the lemniscate will change its shape. When p is less than $c^2/4$, the lemniscate consists of two ovals, one of which contains inside it the point F_1 and the other the point F_2 (Fig. 33). When the product

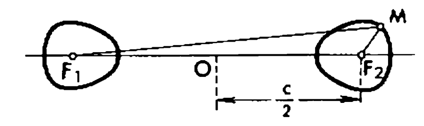


Figure 33

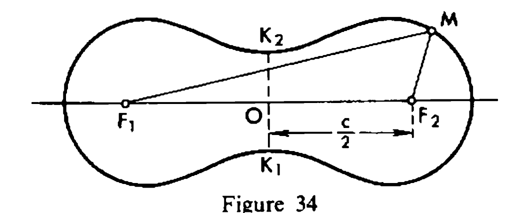


Figure 34

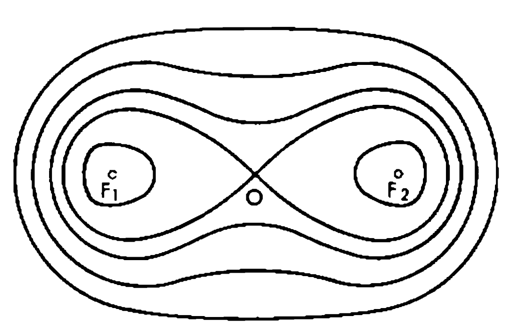


Figure 35

p is greater than $c^2/4$ but less than $c^2/2$, the lemniscate has the form of a biscuit (Fig. 34). If p is close to $c^2/4$, the "waist" of the biscuit K_1K_2 is very narrow and the shape of the curve is very close to the figure-eightshape. If p differs little from $c^2/2$, the waist is hardly noticeable, and for p equal to or greater than $c^2/2$ the waist disappears completely and the lemniscate takes the form of an oval (Fig. 35; the figure also shows some other lemniscates for comparison).

18. The Lemniscate with Arbitrary Number of Foci

Take an arbitrary number of points $F_1, F_2, \ldots, F_n$ in the plane and make a point M move so that the product of its distances from all the points taken is always constant. We shall obtain a curve whose form depends on the relative positions of the points $F_1, F_2, \ldots, F_n$ and the value of the constant product. Such a curve is called a lemniscate with n foci.

We considered above lemniscates with two foci. Using various numbers of foci, placing them in different positions and assigning different values to the product of the distances, one can construct lemniscates with the most peculiar outlines. Let us move the pencil point without lifting it from the paper starting at a point A, so that in the

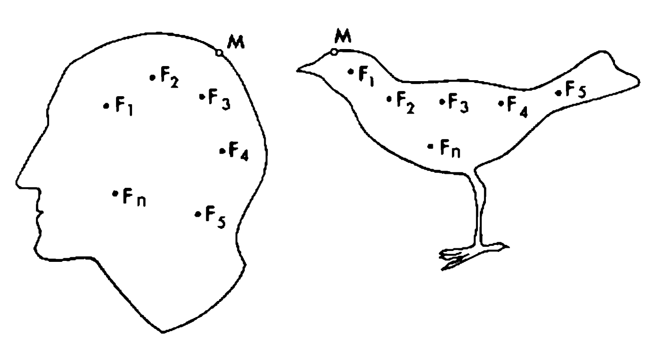


Figure 36

end it returns to the same point A. The pencil will describe a curve; our only condition is that the curve should never intersect itself. It is clear that in this way we can obtain curves with, for instance, the outlines of a human head or a bird (Fig. 36). It turns out that for such an

arbitrary curve it is possible to choose a number n and positions of foci

$$F_1, F_2, \quad F_n$$

and assign a value to the constant product of the distances

$$MF_1 \times MF_2 \times \quad \times MF_n = p$$

such that the corresponding lemniscate will not differ visually from the chosen curve. In other words, the possible deviations of the point M describing the lemniscate from the curve drawn will not exceed the width of the pencil mark (the pencil can be sharpened beforehand as sharply as possible so that the pencil mark is very thin). This amazing fact, indicative of the outstanding variety and richness of forms of the lemniscate with many foci, is proved rigorously in higher mathematics, and the proof is very complicated.

19. The Cycloid

Press a ruler to the lower edge of a blackboard and roll a hoop or a circle (made of cardboard or wood) along it keeping it tight to the ruler and the blackboard. If we fix a piece of chalk to the hoop or circle (at the point of contact with the ruler), the chalk will describe a curve (Fig. 37) which is called a cycloid (which means "circular" in Greek). One revolution of the loop corresponds to one arc of the cycloid $MM'M''N$; if the loop rolls farther, it will generate more and more arcs of the cycloid.

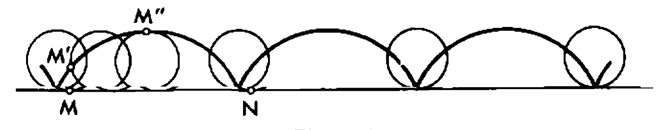


Figure 37

To construct on paper an arc of the cycloid described by the rolling hoop whose diameter is equal, for instance, to three centimetres, mark off on the straight line a segment equal to

$$3 \times 3.14 = 9.42 \text{ cm}$$

Thus, we have a segment the length of which is equal to the length of the rim of the hoop, that is, to the length of the circumference of a circle with diameter three centimetres. Divide this segment into several equal parts (for example, six), and for every point of division draw the

loop in the position when it touches the point (Fig. 38), and number these positions thus

0, 1, 2, 3, 4, 5, 6

In order to go from one position to the next one the loop has to rotate one-sixth of the whole revolution (as the distance between two neighbouring points is equal to one-sixth of the circumference of the circle).

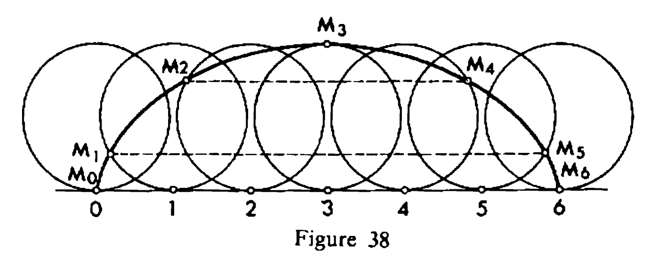


Figure 38

Thus, if in position *0* the chalk is at the point M_0, in position *1* it will be at the point M_1 (at one-sixth of the circumference from the point of tangency), in position *2* at the point M_2 (at two-sixths of the circumference from the point of tangency), and so on. For locating the points M_1, M_2, M_3, etc. we have only to make marks beginning with the point of tangency by circles with a radius equal to 1.5 cm (using compasses), making one mark in position *1*, two consecutive marks in position *2*, three consecutive marks in position *3*, and so on. Now to construct a cycloid there is only one thing left to do: to connect the points

$$M_0, M_1, M_2, M_3, M_4, M_5, M_6$$

by a smooth curve (by hand).

20. The Curve of Fastest Descent

Of the many remarkable properties of the cycloid we shall note only one, for which it was given the sonorous and rather difficult name, the brachistochrone. The name is composed of two Greek words *brachistos* meaning shortest and *chrones* meaning time.

Consider the following question: what form should be given to a well-polished metal trough connecting two given points *A* and *B* (Fig. 39) so that a polished metal ball rolls along this trough from point *A* to point *B* in the shortest possible time? At first it seems that

one should choose a straight trough, as it provides the shortest way from A to B. But we are trying to find the shortest time and not the shortest way, and the time depends not only on the path itself but on the speed of the ball as well. If we bend the trough downwards, the

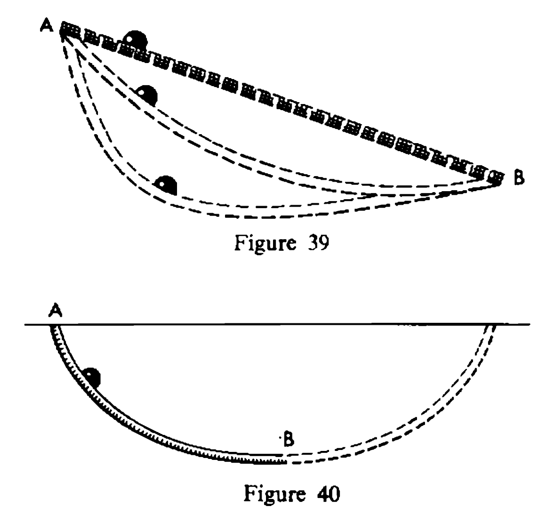


Figure 39

Figure 40

part of it beginning at point A will be steeper than the straight channel, and the ball rolling down such a trough will have a greater speed than on the equidistant segment of the straight trough. But if we make the initial part too steep and comparatively long, then the part near point B will be steeply sloping and also comparatively long; the first part of the path will be covered by the ball very quickly but the second part will be covered very slowly, which may delay the arrival of the ball at point B. Thus, it seems that the trough must be bent downwards but not very steeply.

The Italian astronomer and physicist Galileo Galilei (1564-1642) thought that the path of fastest descent should be in the form of an arc of a circumference. But the Swiss mathematicians Jakob and Johann Bernoulli (brothers) proved by exact calculations that it was not so and that the trough should be bent in the form of an arc of a cycloid (turned upside down, Fig. 40). Since that time the cycloid has also been called the brachistochrone, and the proof of the Bernoullis laid the foundation of a new branch of mathematics, the calculus of variations. This subject deals with determining the forms of curves for which some quantity or other in which we are interested reaches its minimum (or in some cases, its maximum) value.

21. The Spiral of Archimedes

Imagine an infinitely long second hand on a clock, along which a small beetle runs indefatigably at a speed of v cm/sec, starting at the centre of the clock face. After one minute the beetle will be at a distance of $60v$ cm from the centre, in two minutes $120v$ cm and so on. In general, t seconds after the beginning of the race the beetle will be at a distance of vt cm from the centre. During this period of time the second hand will rotate about an angle of $6t°$ (the hand rotates through an angle of $360°/60 = 6°$ every second). Therefore, the position of the beetle on the face of the clock after any number of seconds t is determined in the following way. Mark off the angle α containing $6t°$ from the initial position of the hand in the direction of its rotation and measure off the distance $r = vt$ cm along the new position of the hand. We shall find the beetle in this position (Fig. 41).

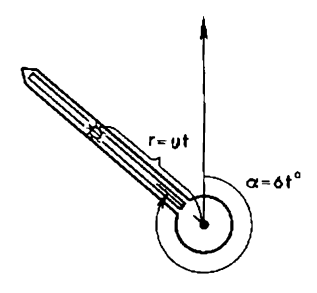


Figure 41

Evidently the relation between the angle of rotation α of the hand (in degrees) and the distance covered r (in centimetres) will have the form

$$r = (v/6)\alpha$$

In other words r is directly proportional to α, the factor of proportionality being equal to $k = v/6$.

Let us adjust a tiny but inexhaustible bottle of black ink to our runner and assume that the paint, flowing out from a tiny hole, leaves on the paper (pasted on to the second hand) a trace of the beetle carried away by the hand of the clock. Then gradually a curve is traced out. This curve was first studied by Archimedes (287?-212 B.C.). In his honour it is called the spiral of Archimedes. It should be noted that Archimedes spoke neither about a second hand nor a beetle (at that time there were not even mechanical clocks with springs: they were in-

vented in the seventeenth century). We introduced them here just for the sake of clarity.

The spiral of Archimedes consists of an infinite number of coils. It starts at the centre of the clock face and goes away farther and farther from it as the number of coils increases. In Fig. 42 you can see the first coil and part of the second.

You might have heard that it is impossible to divide an arbitrary angle into three equal parts using only a straightedge and compasses (in particular cases, when an angle contains for instance 180°, 135° or 90°, this problem is easily solved). But if we use an accurately drawn spiral of Archimedes we can divide any angle into an arbitrary number of parts.

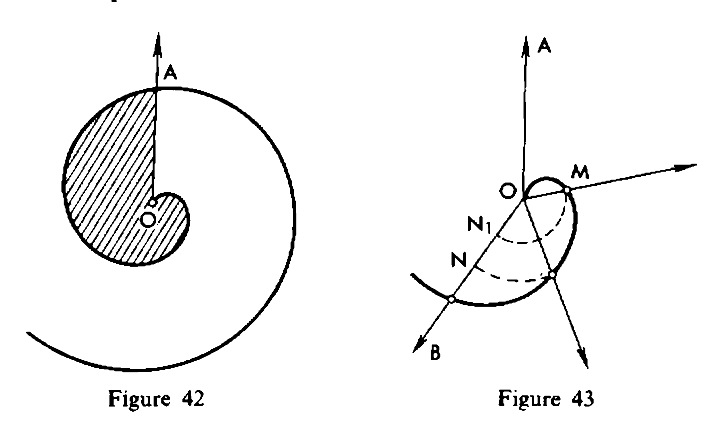


Figure 42 Figure 43

Let us divide, for example, the angle *AOB* into three equal parts (Fig. 43). If we assume that the hand has just rotated through this angle, the beetle is at the point *N* on the side of the angle. But when the angle of rotation was one-third of *AOB*, the beetle was one-third of the distance from the centre *O*. In order to find this position of the beetle we have to divide the segment *ON* into three equal parts. This can be done by using a straightedge and compasses. We get the segment ON_1 whose length is one-third of *ON*. But if we want the beetle to be on the spiral we have to make a cut with circle whose radius is ON_1 (again with the help of compasses!), getting the point *M*. The angle *AOM* will be one-third of the angle *AON*.

22. Two Problems of Archimedes

Archimedes himselfs was interested, however, in more difficult problems, which he himself formulated and solved. These problems are:

(1) determine the area of the figure bounded by the first coil of the spiral (in Fig. 42 it is shaded);

(2) find how to construct the tangent to the spiral at a point N.

It is of great interest that both problems are the earliest examples of problems in mathematical analysis. Since the seventeenth century areas of figures have been calculated by mathematicians with the help of integrals, and tangents have been constructed with the help of derivatives. Hence Archimedes might be called the precursor of mathematical analysis.

As far as the first problem is concerned we shall simply cite the result obtained by Archimedes: the area of the figure is exactly equal to one-third of the area of a circle with radius OA. But for the second problem we can indicate a method of solution. It is a simplified version of Archimedes' own argument. The argument depends on the fact that the velocity of the beetle describing a spiral at any point N is directed along the tangent to the spiral at this point. If we know the direction of the velocity, we are able to construct the tangent.

The beetle's movement at the point N is composed of two different movements (Fig. 44): one in the direction of the second hand at a speed of v cm/sec, the other along the circumference of the circle with centre O and radius ON. To visualise the latter suppose that the

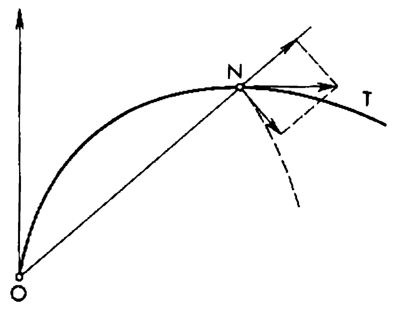


Figure 44

beetle stood stockstill for a moment at point N. Then it would be carried away by the second hand along the circumference of the circle of radius ON. The velocity of the circular movement is directed along the tangent to the circle. And what is its magnitude? If the beetle could go around the whole circumference with the radius ON, then in 60 seconds he would cover a distance equal to $2\pi \times ON$ cm. In this case the speed would remain constant, so to determine it we have to divide the distance by the time. In this way we get $2\pi \times ON/60 = \pi \times$

$\times ON/30$ cm/sec, which is a little more than $0.1 \times ON$ cm/sec ($\pi/30 \approx 3.14/30 \approx 0.105$).

Now we know both components of the velocity at the point N: one in the direction of ON and equal to v cm/sec, the other perpendicular to it and equal to $\pi \times ON/30$ cm/sec. All we have to do is to add them according to the parallelogram law. The diagonal determines the velocity of the combined motion and at the same time shows the direction of the tangent NT to the spiral at any given point.

23. The Chain of Galilei

In his book "Dialogues on Two New Sciences", which was first published in Italian in Leiden (the Netherlands) in 1638, Galilei suggested the following method of constructing a parabola: "Drive two nails into a wall at a convenient height and at the same level; make the distance between them twice the width of the rectangle upon which it is desired to trace the semiparabola. Hang between these two nails a light chain of such a length that the depth of its sag is equal to the height of the rectangle (Fig. 45). The chain will then assume the form

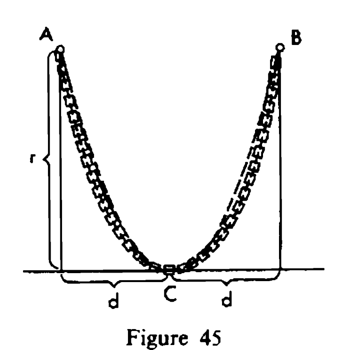


Figure 45

of a parabola, so that if this form is marked by points on the wall, we shall have described a complete parabola, which can be divided into two equal parts by drawing a vertical line through a point midway between the two nails".

The method is simple and graphic but not exact. Galilei understood this himself. In fact, if we draw a perfect parabola, there will be a gap between the real parabola and the chain. One can see this in Fig. 45, where the corresponding parabola is a dotted curve.

24. The Catenary

Only half a century after the publication of Galilei's book the elder of the Bernoulli brothers, Jacob, found in a purely theoretical way an exact formula giving the shape of a suspended chain. Not publishing his findings he challenged other mathematicians to do what he had done. That was in 1690. In 1691 the correct solution was published by Jakob Bernoulli himself and also by Christian Huyghens, Gottfried Wilhelm von Leibniz and the younger brother of Jakob, Johann Bernoulli. All of them used in the solution of the problem the laws of mechanics and the powerful technique of the recently developed mathematical analysis, the derivative and integral.

The curve in which a chain hangs when suspended from two points was called a catenary by Huyghens.

As there are chains of different length and the points of suspension can be at different distances from each other, there exists not one but many catenaries. But all of them are similar in the same way that all circles are similar.

25. The Graph of the Exponential Function

It turned out that the clue to the secret of the catenary was hidden in the exponential function. In the eighteenth century it was a novelty but now it should be known to every high school student. This func-

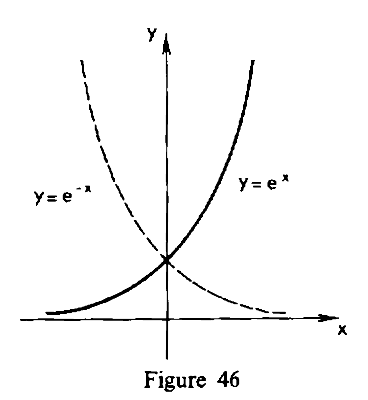


Figure 46

tion has the form $y = a^x$, where a is a positive integer distinct from unity. It became clear from calculations that the most convenient way for drawing a catenary was to set a equal to e, the base of natural

logarithms. These are sometimes called Napeirian logarithms after the Scottish mathematician John Napier (1550-1617), one of the discoverers of logarithms. Number e is almost as famous as the number π; its approximate value calculated with an accuracy of 0.0005 is $e \approx 2.718$.

In Fig. 46 you can see two graphs: the solid line depicts the graph of the exponential function $y = e^x$, and the dotted line the graph of another exponential function closely connected with the previous one: $y = (1/e)^x (1/e \approx 0.368)$.

If we use negative powers, the latter function may be represented in the form $y = e^{-x}$ Now it is clear that both graphs are mutually symmetric with respect to the y-axis, which can be seen from the figure.

Now we shall form two new functions, taking for every x the half-sum of our exponential functions, $y = (1/2)(e^x + e^{-x})$, or the half-difference $y = (1/2)(e^x - e^{-x})$. The graphs of these two new functions are shown in Figs. 47 and 48. It turns out that the first graph is a catenary. By simple transformations which will be discussed below we can obtain any catenary symmetric with respect to the y-axis. As far as the graph in Fig. 48 is concerned we shall use it to help in the transfer from the catenary of Fig. 47 to the more general case.

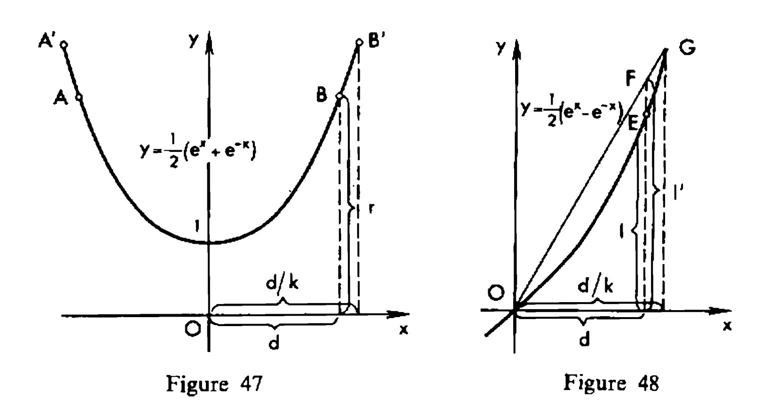


Figure 47

Figure 48

26. Choosing the Length of the Chain

Let us consider in more detail the relation between the curve in Fig. 47 and the form of the suspended chain. Imagine that the curve is drawn on a perfectly vertical, polished wall and that we can drive in

nails at different points of the curve. Let us drive them in at points A and B that are at the same level, as Galilei suggested (this condition is not essential, however). Choose a light chain whose length is exactly $2l$, the length of the arc AB, and fix its ends at A and B. The chain will hang in the form of the arc shown above. There will be no gaps between the hanging chain and the curve.

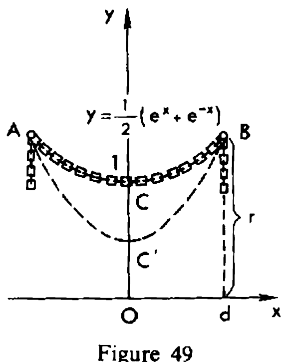


Figure 49

The selection of a chain of the required length can be done by trial and error. We can take a longer chain and hang it by different links at the points A and B increasing or decreasing, as needed, the suspended part till there is coincidence (Fig. 49). But there is another way: knowing d (half the distance between the nails), one can calculate l (half the length of the arc AB) and take a chain whose length is $2l$. This calculation can be done using integrals. We shall give only the result: $l = (1/2)(e^d - e^{-d})$. It follows that if we take $x = d$ in the graph of the function $y = (1/2)(e^x - e^{-x})$ (Fig. 48), the corresponding value of y of the point E will be l. Since $l = (1/2)(e^d - e^{-d}) < r = (1/2)(e^d + e^{-d})$ (Fig. 49), we can make a rather curious conclusion: the length of the arc CB of the catenary in Fig. 49 (half the length of the chain) is shorter than the ordinate of the point of suspension. On the other hand $l > d$, that is, this length is greater than the abscissa of the point of suspension.

27. And What if the Length is Different?

How can we deduce the equation of the catenary for the case when for given points of suspension A and B the length of the chain, $2l'$, is not equal to the length $2l$ of the arc AB of the curve $y =$

$= (1/2)(e^x + e^{-x})$? In looking for the answer we shall depend on the fact mentioned above that all catenaries are similar.

Suppose $l' > l$. Then the chain will hang along some arc $AC'B$ beneath the arc ACB (Fig. 49). We shall show that the required equation of the catenary to which the arc $AC'B$ belongs can be found in two steps. First we have to transform the curve *1*, $y = (1/2)(e^x + e^{-x})$, into the curve 2, $y=(k/2)(e^{x/k}+e^{-x/k})$. The latter can be obtained from curve *1* by the similarity transformation with the point O as centre and k as the ratio of magnification (k is positive). Then we have to transform the curve *2* into the curve *3*, $y=b+(k/2)(e^{x/k}+e^{-x/k})$, by shifting the former in the direction of the x-axis (upwards or downwards, depending on whether $b>0$ or $b<0$).

The trick is to determine the ratio of magnification. For this purpose mark the point F with the coordinates $x=d$ and $y=l'$ in the plane of the auxiliary curve depicted in Fig. 48. Since $l' > l$, it will not be on the curve but above it.

Extend OF till it intersects the curve at some point G (it can be proved that apart from point O there is just one point of intersection). Assume $OF/OG=k$ (in our case $0<k<1$); then the coordinates of the point G will be $x = d/k$ and $y = l'/k$ (prove this!). Hence they will be related by the equation of the curve: $l'/k=(1/2)(e^{d/k}-e^{-d/k})$. From this it follows that if we take the points A' and B' with the abscissas $-d/k$ and d/k on the curve *1* (Fig. 47), the length of the arc $A'B'$ connecting them will be equal to $2l'/k$ (see Section 26).

28. All Catenaries are Similar

We shall use the number k obtained as the ratio of magnification in the transformation of curve *1*; for the centre of similarity we take the origin of coordinates O. Then to each point $P(x, y)$ of curve *1* there will correspond a point $Q(kx, ky)$ of the transformed curve *2* (Fig. 50). If we set $X = kx$ and $Y = ky$, we get $x = X/k$ and $y = Y/k$. The latter two numbers must satisfy the equation for curve *1*, as the point $P(x, y)$ is on this curve: $Y/k = (1/2)(e^{X/k} + e^{-X/k})$. But this is exactly the equation for curve *2* obtained as the result of the transformation. The capital letters denoting the coordinates can be replaced by lower case, bearing in mind that now they represent the coordinates of an arbitrary point on curve *2*.

Note that the points A' and B' of the curve *1* with abscissas $-d/k$ and $+d/k$ correspond to the points A'' and B'' of the curve *2* with the abscissas $-d$ and $+d$ (Fig. 51). By virtue of the similarity of the arcs $A'B'$ and $A''B''$ it follows that the length of $A''B''$ is equal to $(2l'/k)k = 2l'$, that is, to the given length of the chain. This is the advantage of

the curve *2* over the initial curve *1*. Its drawback, however, lies in the fact that the curve *1* passes through the points of suspension A and B, and the curve *2* may not pass through them. But this drawback can be eliminated easily. If the ordinate of the point B'' (or A''), $(k/2)\,(e^{d/k} +$

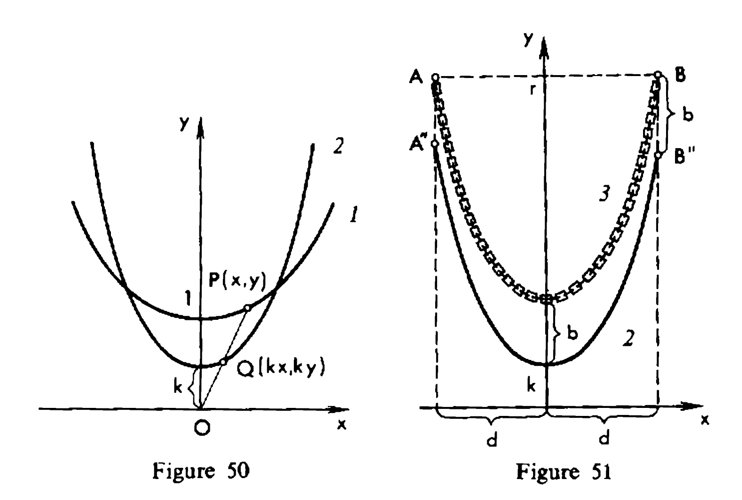


Figure 50

Figure 51

$+\, e^{-d/k})$, is not equal to r, that is, B'' does not coincide with B, then we set $r - (k/2)(e^{d/k} + e^{-d/k}) = b$.

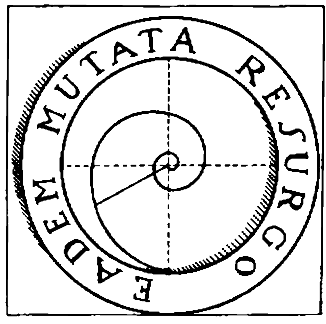
As the result of the shift of curve *2* by a quantity b in the direction of the x-axis it will be transformed into curve *3*: $y = b + (k/2)(e^{d/k} + e^{-d/k})$. The latter curve is, in the first place, similar to curve *1* and is consequently a catenary. In the second place, it passes through the given points of suspension: $A = (-d, r)$ and $B(d, r)$. And in the third place, the length of the arc AB is equal to that of the chain $2l'$. These conditions are sufficient for the curve to hang along the arc AB, as was proved by Bernoulli, Huyghens and Leibniz.

29. The Logarithmic Spiral

This curve might well have been named after Descartes as he was the first to mention it in one of his letters (1638). However, its detailed investigation was made by Jakob Bernoulli only half a century later. Contemporary mathematicians were impressed by its properties. The stone tomb on the grave of Bernoulli is decorated with the coils of the logarithmic spiral (Fig. 52).

We have seen already (Section 21) that the spiral of Archimedes is generated by a point moving along a ray (the "infinite hand") in a way

such that the distance from the origin of the ray increases proportionally to the angle of its rotation: $r = k\alpha$. The logarithmic spiral can be obtained if we demand that not the distance itself but its logarithm should increase in direct proportion to the angle of rotation. Usually



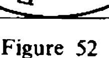


Figure 52

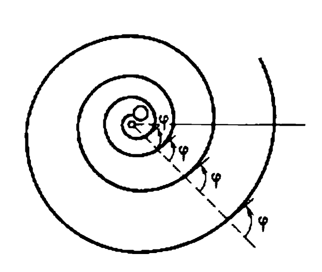


Figure 53

the equation of the logarithmic spiral is written using the number e as the base of natural logarithms (Section 25). Such a logarithm is denoted by $\ln r$. Thus, the equation of the logarithmic spiral is written in the form

$$\ln r = k\alpha$$

or, which is the same,

$$r = e^{k\alpha}$$

Nothing prevents us from continuing to measure the angle of rotation α in degrees, but mathematicians prefer to measure it in radians, that is, to use the ratio of the length of the arc of a circle between two sides of the central angle to the radius of the circle. Then the rotation of the hand through a right angle is measured by the number $\pi/2 \approx 1.57$, the rotation through a straight angle is measured by the number $\pi \approx 3.14$ and a full rotation, measured in degrees by the number 360, by the number $2\pi \approx 6.28$.

We shall mention only one of many properties of the logarithmic spiral, namely, that any ray issuing from the centre intersects every coil of the spiral at the same angle. The magnitude of the angle depends only on the number k in the equation of the spiral. The angle between the ray and the spiral is understood as the angle between the ray and a tangent to the spiral drawn at the point of intersection (Fig. 53).

30. The Involute of a Circle

Imagine a goat grazing in a meadow. It is tied up to a stake whose cross section is a circle with a long rope. Stretching the rope taut, the goat browses on the grass without noticing that the rope, winding around the stake, becomes shorter and shorter. At last the goat is drawn tight to the stake. It cannot guess that to get out of its difficulty

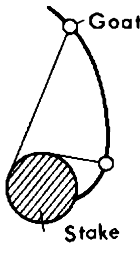


Figure 54

it should go in the opposite direction so that the rope begins to unwind. What curve is described by the goat in this case? In order not to occupy ourselves with drawing a goat we have replaced it by a small circle in Fig. 54. You may consider it to be a schematic drawing of a collar to which the rope is fixed. You may know that the arc along which the goat is moving away from the stake (earlier it was approaching the stake along the same arc) belongs to an infinite curve which is called the involute of a circle. Mathematicians got acquainted with this curve for the first time in the eighteenth century. The French

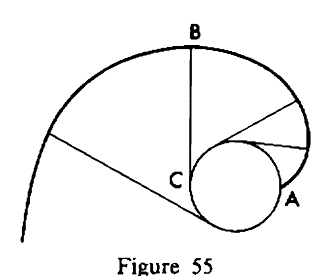


Figure 55

Figure 56

philosopher and writer Denis Diderot (1713-1784) studied its properties in 1748.

We shall now give a more precise definition of the involute of a circle. Imagine that an infinitely long, thin and inextensible string is

wound on a circle so that one of its ends remains free. Then we shall tie a pencil to it. If we begin to unwind the string keeping it taut, then the point of the pencil will describe a spiral-shaped curve which is called the involute of a circle (Fig. 55). It follows from the definition that for any point B on the curve the length of the free part of the string BC is exactly equal to the length of the arc of the circle, AC.

If we unwind the string from another curve distinct from a circle, for instance from an ellipse, then we shall have the involute of an ellipse (Fig. 56).

Conclusion

We are finishing our outline sketch of some remarkable curves. We have considered here only a few of them, and we have come nowhere near exhausting their properties. The aim of the author was to get the reader, already familiar with elementary mathematics, interested by some curious facts from the inexhaustible treasure-house of mathematical knowledge. In doing this proofs and calculations were omitted.

If we draw a comparison with a trip to the zoo, we could say that the author showed the reader a peculiar "collection of curves", spending little time at every cage to get acquainted with a curve and confining himself with a very simple characterization of its "habits" The reader wishing to enrich his knowledge in this field should read the book "*What Is Mathematics?*" by R. Courant and H. Robbins (1948). Another useful book is *Handbook of Curves and Their Properties* by Robert C. Yates (1952). The reader with a little knowledge of Russian can find interesting articles on curves and their properties in issues of the magazine *Kvant* (The Quantum). It should be noted that a more detailed study of curves is impossible without a deeper mathematical knowledge, and, in particular, without an acquaintance with integral and differential calculus.

TO THE READER

Mir Publishers welcome your comments on the content, translation and design of this book.

We would also be pleased to receive any proposals you care to make about our publications.

Our address is:

USSR, 129820, Moscow I-110, GSP
Pervy Rizhsky Pereulok, 2
Mir Publishers